Early Chronic Illness Detection
Group 27

VICKEY
2021299

SUYASH
2021293

ADITYA
2021305

SATYAM
2021285

SATYAM PANDEY
2022463

September 8, 2024

1 Motivation

The motivation for this project stems from the growing need to address mental health issues through data-driven approaches. Depression affects millions worldwide, but early detection and intervention remain challenging. By leveraging machine learning techniques, this project aims to predict the likelihood of depression based on various health, lifestyle, and socio-economic factors. Predicting depression could lead to earlier intervention, potentially reducing the severity of the condition and improving quality of life. This model may offer valuable insights for mental health professionals, enabling them to identify high-risk individuals and tailor treatments more effectively. The idea for this project was inspired by research on the predictive power of lifestyle and socio-economic factors in mental health. We aim to build upon this research and develop a model that provides accurate and reliable predictions using the dataset we have chosen

2 Related Work

- Previous studies have employed machine learning models to predict mental health outcomes, such as depression, using health and lifestyle data. One study used decision trees and support vector machines to classify individuals based on their risk of depression, achieving promising results.

- Another research project explored the predictive power of socio-economic status, finding significant correlations between income, education, and mental health. Their use of regression models demonstrated the feasibility of predicting depression based on these features.[]

- Other works have focused on incorporating additional health metrics, such as sleep patterns and chronic medical conditions, into predictive models for mental health, demonstrating the importance of holistic data.

3 Expected Timeline

- **Week 1** Dataset cleaning and exploration
- **Week 2-3** Pre-processing of the dataset and feature engineering.
- **Week 4** Exploratory data analysis, visualization, and feature selection.
- **Week 5-6** Model selection and baseline model creation.
- **Week 7-8** Training and validation of machine learning models.
- **Week 9** Model evaluation and fine tuning.

4 Individual Tasks

- Dataset exploration and cleaning, preprocessing dataset, and feature engineering - Vickey Kumar, Suyash
- Exploratory data analysis and visualization, model selection and baseline model setup - Aditya, Satyam Pandey
- Hyper-parameter tuning and validation - Satyam
- Model evaluation and final adjustments - Everyone
- Documentation and presentation of findings - Everyone

5 Final Outcome

- The project will deliver a machine learning model to predict depression risk based on health, lifestyle, and socio-economic factors.

- The model will help mental health professionals identify individuals at risk for early intervention.

- Multiple rounds of training, tuning, and validation will ensure the model is accurate and reliable.

- Key insights, including visualizations, will highlight relationships between various factors and depression risk.

- The ultimate goal is to support research and practical applications in mental health by offering data-driven insights into depression prediction.